Near-Infrared Fusion for Photorealistic Image Dehazing

Frederike Dümbgen*, Majed El Helou*, Natalija Gucevska, Sabine Süsstrunk
School of Computer and Communication Sciences, EPFL, Lausanne, Switzerland.
****equal contribution***

Abstract

Scattering of light due to the presence of aerosol particles along the path of radiation causes atmospheric haze in images. This scattering is significantly less severe in longer wavelength bands than in shorter ones, thus the importance of near-infrared (NIR) information for dehazing color images.

This paper first presents an adaptive hyperspectral algorithm that analyzes intensity inconsistencies across spectral bands. It then leverages the algorithm's results to preserve photorealism of the visible color image during the dehazing. The color images are dehazed through a hyperspectral fusion of color and NIR images, taking into account any inconsistencies that can affect the photorealism. Our dehazing results on real images contain no halo or aliasing artifacts in hazy regions and successfully preserve the color image elsewhere.

Keywords: *Image dehazing, joint color-NIR imaging, Rayleigh scattering, hyperspectral fusion, photorealism.*

Introduction

Camera sensors are made of silicon that is sensitive to the part of the electromagnetic spectrum ranging from visible light (blue, green and red) to NIR. In most digital cameras, a filter is placed in front of the sensor to block the NIR radiation and stop it from affecting the acquisition of the color image [1]. Using recent research [2, 3, 4, 5], a camera can simultaneously capture both visible and NIR images after removing the NIR-blocking filter. Image dehazing [6] is one of many applications where the additional NIR image proves to be very useful. Scene recognition [7] or more specifically shadow detection [8] and face recognition under different lighting conditions [9] are some of the multiple other applications.

NIR radiation is the part of the electromagnetic spectrum with wavelengths ranging approximately from 700 to 1100 nm. Therefore, it has larger wavelength compared to visible light which ranges in wavelength from 400 to 700 nm. According to Rayleigh scattering [10], as long as haze particles are smaller than one tenth of the radiation wavelength, the scattered energy E_s is inversely proportional to the wavelength to the fourth power,

$$E_s \propto \frac{E_i}{\lambda^4} \tag{1}$$

where λ is the radiation wavelength and E_i is the energy of incident radiation. Having a larger wavelength than visible light, the energy lost due to scattering in the NIR band is lower compared to the visible band. NIR radiation thus has an advantage over visible light with respect to haze and it has first been leveraged for dehazing by Schaul *et al.* [6].

Although visible light and NIR are correlated, mostly through their high-frequency components, materials have different reflective properties in NIR compared to visible light [11]. These differences or inconsistencies, most apparent in low-frequency regions, have to be taken into consideration during hyperspectral fusion when the output needs to be photorealistic. Throughout this paper, a result is called photorealistic if it is close to the visual perception of the real-world scene, i.e. if the results' colors are in accordance with the original color image.

Figure 1. *Image dehazing results obtained using the proposed algorithm. From top left to bottom right are shown the hazy input image, NIR image, inconsistency mask, and the dehazed image. Note how the details of the NIR image are merged with the RGB image without changing the color of non-hazy regions such as the port in the bottom of the image.*

In this paper, we present a novel adaptive algorithm that detects intensity inconsistencies between color and NIR in a captured scene and improves on state-of-the-art dehazing results. By leveraging the inconsistency mask, no artifacts are carried from NIR to color during the dehazing process. Our proposed solution improves upon the state-of-the-art multi-scale method developed in [6] by adjusting the color-NIR fusion at every scale based on our inconsistency mask, thus successfully eliminating the artifacts of [6]. Figure 1 shows an example of a hazy color image and its NIR counterpart in the first row, and our inconsistency mask and dehazed color image in the second row.

We first present a survey of the state-of-the-art approaches for image dehazing. Our solution is detailed in Section *Proposed Solution*. Finally, the improvement of our algorithm over the state of the art is highlighted through experimental results in Section *Experimental Results*.

Literature Review

The classical dehazing approach consists of haze detection followed by haze removal. Haze removal is typically performed

by restoring constrast and detail in the hazy regions of the image. On one hand, this can be achieved using intensity redistribution. Some of the oldest methods use first and second order image statistics [12], while newer approaches use histogram equalization conserving brightness [13] or image luminance [14], respectively, or adaptive histogram equalization avoiding the loss of detail and artifacts [15]. These methods perform particularly well when the images exhibit a poor intensity distribution. On the other hand, tools such as wavelet [16], ridgelet [17] and curvelet [18] transforms perform edge-enhancement rather than intensity redistribution, also leading to the desirable effect of increasing contrast in hazy regions.

Despite its difficulty, some authors have proposed single-image dehazing solutions without using additional information. Tan [19] proposes a single-image method which maximizes local contrast under the condition that airlight is smooth. Fattal *et al.* [20] exploit the physical properties of the acquisition, in a similar way as the polarization filters used at acquisition time by some photographers [21]. He *et al.* [22] notice that in many natural scenes very dark pixels exist in every local image patch, and they use this observation to both obtain a depth map based on the attenuation of dark pixels and to dehaze the image accordingly. Haze detection is facilitated when additional information is available. It can consist of multiple images in different weather conditions [23], interactive inputs provided by a user [24], 3D models provided by additional depth sensors or multiple views [25], or a library of georeferenced digital terrain and urban models [26].

Because of the robustness of NIR to haze, it provides a nearly haze-free version of the scene, which can be exploited for dehazing the color image. The main challenge in this appraoch is to integrate hyperspectral information without distorting the visible information, i.e. to conserve photorealism. Schaul *et al.* [6] obtain state-of-the-art results by using the edge-preserving filter proposed by Farbman *et al.* [27] to merge details from the NIR channel into the RGB channel. Their results show excellent dehazing performance but lack in photorealism: the dehazed images have artificially high intensities in regions of high NIR reflectance, such as vegetation. Feng *et al.* [28] leverage NIR information for dehazing by estimating the airlight. With airlight estimation and a simplified haze model, the haze-free image can be attained through an optimization. Even though only a suboptimal solution can be reached, good dehazing is obtained. However, due to inaccurate airlight estimation, the overall color of the image is affected, making it often darker and more saturated than the original.

More general hyperspectral fusion algorithms can be applied to merge the NIR information into the visible image. Connah *et al.* [29] propose a spectral mapping algorithm from a general hyperspectral image to a subspace of lower dimension. While this approach is designed to yield relatively photrealistic fusion results, it is suboptimal for the dehazing problem. For dehazing, it is more important to fuse the details in hazy regions from NIR into the visible image, but Connah's approach aims to keep the most information from all channels. Haze in the RGB image being seen as information it is thus conserved during the fusion, instead of being corrected based on higher-frequency NIR information.

Proposed Solution

Healthy vegetation generally appears brighter in the NIR channel compared to the visible color band. Note for example the difference in intensity of the patch of grass in Figure 1. This is because healthy plants do not absorb NIR. Indeed, NIR radiation travels inside the leaves to only interact with mesophyll cells that play a role in plant photosynthesis. Half of the NIR waves get reflected and half cross the leaf unaffected. However, healthy vegetation is not the only source of intensity inconsistencies between NIR and visible color. Different materials can cause such inconsistencies depending on their reflectance properties. Our first goal is to detect all such inconsistent pixels in the captured scene.

Due to spectral proximity, NIR is most correlated to red among all color channels [11]. Hence, if an inconsistency is present in the NIR relative to the red channel, it is very likely present relative to all other less correlated color bands. We also experimentally find that among red, green, blue, luma, and luminance channels, the most robust channel for inconsistency analysis is the red channel and it is used throughout this paper. In fact, comparing NIR to the red channel for vegetation detection is on par with the remote sensing community [30].

We propose a statistical adaptive method for the detection algorithm. First, the difference image Δ is computed pixel-wise,

$$\Delta(x,y) = NIR(x,y) - Red(x,y). \tag{2}$$

We then make a rough estimate of the percentage of pixels having intensity inconsistency,

$$p = \frac{100}{m.n} \sum_{x=1}^{m} \sum_{y=1}^{n} u(\Delta(x,y) - t_0), \tag{3}$$

where the image is mxn pixels, t_0 is a fixed threshold and $u(\cdot)$ is the Heaviside function. We continue by standardizing the difference image with the z-score operation,

$$\Delta_Z(x,y) = \frac{\Delta(x,y) - \bar{\Delta}}{\sqrt{\frac{1}{m.n}\sum_{x=1}^{m}\sum_{y=1}^{n}(\Delta(x,y) - \bar{\Delta})^2}}, \tag{4}$$

where $\bar{\Delta}$ is defined as the average of Δ, which yields a zero-mean and unit-variance difference map. Assuming all scene images have the same probability distribution for a given pixel to be part of inconsistencies (up to mean and variance shifts), then a fixed threshold would result in a constant known probabilistic threshold for considering a pixel to be representing an intensity inconsistency.

However, the above assumption doesn't apply to all images (note that even among vegetation there are variations in the Δ map). This means that classifying as inconsistency the $X\%$ most likely pixels could yield the exact solution when $X\%$ of the image is made up of inconsistencies to begin with. If the scene contains no inconsistencies or is only made up of inconsistencies or any percentage different from X, the classification would be erroneous. This is where we leverage our *a priori* estimate p from (3) to linearly adapt our probabilistic threshold on the set $[0, 1.5]$,

$$t_{0Z} = 1.5 - (\frac{p}{100}) * 1.5. \tag{5}$$

A fixed range such as $[0, 1.5]$ can be used since our map is always normalized to zero mean and unit variance. A very low *a priori*

percentage of inconsistencies is mapped to a threshold close to 1.5, meaning that differences need to be above 1.5 standard deviations away from the mean to be considered true positives. On the other hand, a high *a priori* percentage of inconsistencies, is mapped to values close to 0. Lastly, the inconsistency mask is obtained as

$$M(x,y) = u(\Delta_Z(x,y) - t_{0Z}) \tag{6}$$

and the probability for each pixel to be inconsistent is given by

$$P(x,y) = M(x,y) \cdot \left[(1-\alpha) + \alpha \cdot \frac{(\Delta_Z(x,y) - t_{0z})}{\max(\Delta_Z(x,y) - t_{0z})} \right], \tag{7}$$

where $\max(\cdot)$ returns the maximum value of the map $(\Delta(x,y) - t_{0z})$. Equation (7) follows the *soft thresholding* LASSO estimator [31] and bounds $P(x,y)$ to the interval $[1-\alpha, 1]$ where α is a real parameter in $[0,1]$.

Schaul *et al.* [6] use an edge-preserving filter to decompose both color and NIR images into K levels of detail and coarseness. The operation is carried on the luma channel of the color image. The multi-scale detail images of the color luma and the NIR channel are combined back while keeping the detail pixels with higher intensity. With this operation, the additional high frequency present in the NIR image is artificially added to hazy color regions. The problem is that vegetation and many other materials, being naturally brighter in NIR, are carried into the color image and distort it. By incorporating the probability estimate into the multi-scale fusion step, a photorealistic dehazed image can be obtained,

$$F_0(x,y) = V_K^a \prod_{k=1}^{K} \max(V_k^d(x,y), N_k^d(x,y) * (1 - P(x,y))) + 1, \tag{8}$$

where F_0 is the fused luma, V_k^d and N_k^d are the detail images of the visible channel luma and the NIR at decomposition level k, respectively, and V_K^a is the K-th approximation layer. The role of the factor $(1 - P(x,y))$ in (8) is to attenuate the NIR detail image where high intensity is due to highly reflective materials in NIR and not haze in visible color. The fused luma F_0 is finally recombined with the blue-difference and red-difference chroma components extracted from the original color image to obtain the final dehazed color image.

Experimental Results

We test the proposed approach on images captured using a Canon Rebel T1i camera [6]. The built-in infrared-blocking filter of the camera sensor is removed. Color images are obtained by placing an external infrared-blocking filter in front of the lens before capture. Likewise, a filter is mounted on the camera lens to block visible light during the NIR image acquisition. Therefore, every scene is captured twice. However, the scenes being motion-free, the two images match perfectly.

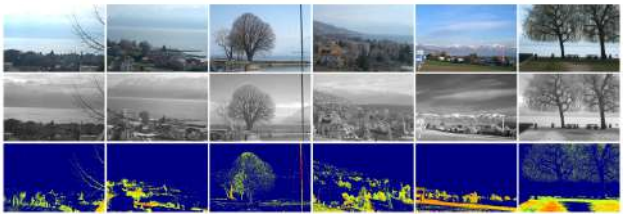
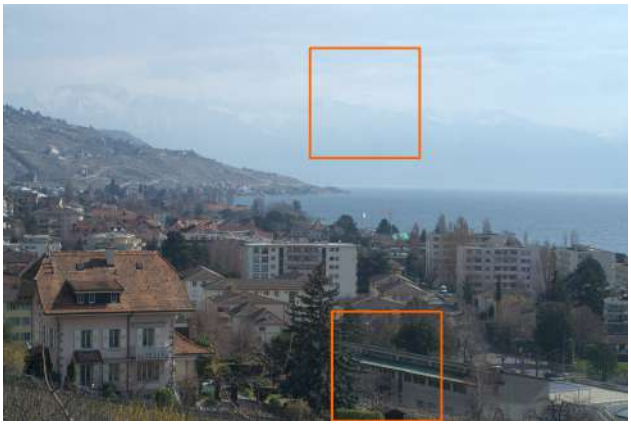


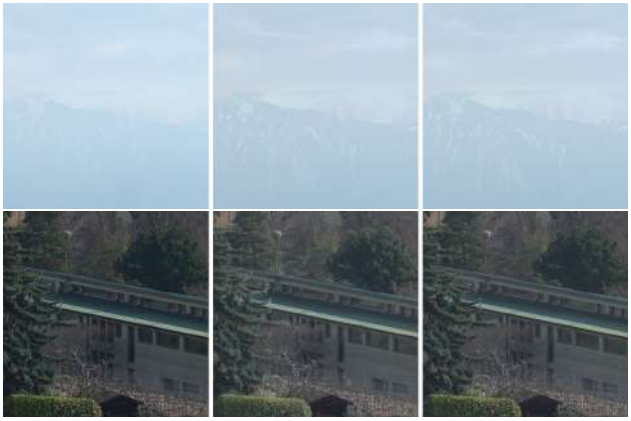
Figure 2. *Dehazing masks (bottom) obtained by comparing the NIR information (middle) with the visible images (top) for images from the dataset provided by [6].*

The dehazing masks obtained on the dataset provided by [6] are shown in Figure 2. The masks reliably depict the areas with high NIR responses due to vegetation, and also detect outliers in NIR intensity resulting from other materials such as reflecting street signs. Note also that not all vegetation has a high NIR response; the intensity depends among other factors on the healthiness of the plant. We compare our approach to the state-of-the-art dehazing pipeline proposed in [6] (see Figure 3). Without the inconsistency mask, vegetation-rich regions appear artificially bright in the dehazed image, which can lead to illusions of snow or dust on the plants. The two crops of Figure 3 (b) highlight the advantages of our proposed algorithm. The trees keep their natural dark color, while the dehazing performance is identical to [6]. The same conclusion can be drawn from looking at the results shown in Figure 4.

Our algorithm's performance is compared to other results on



(a) *Original hazy color image*



(b) *Original image (left) compared to dehazed image by Schaul [6] (middle) and our result (right).*

Figure 3. *Image dehazed using the algorithm of Schaul [6] and the proposed solution. The regions with vegetation exhibit too bright responses in the result by Schaul due to high NIR intensities. By leveraging the dehazing mask, this effect is attenuated in our approach.*

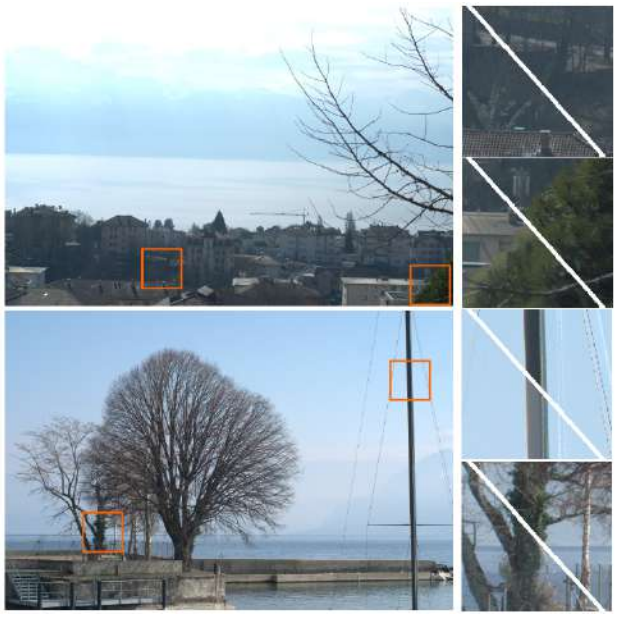


Figure 4. *Original hazy images (left) and crops of results obtained after dehazing (right) using our algorithm (top) and the algorithm proposed by Schaul* et al. *[6]. Artificially bright regions resulting from high material reflectance properties are corrected using our algorithm.*

two images in Figure 5. The algorithm by He *et al.* [22] dehazes effectively but changes the intensity distribution of the original image significantly and yields artificial looking results. While the algorithm of Feng *et al.* [28] delivers very good dehazing performance, it significantly affects the overall brightness and saturation, yielding results far from the visually perceived image. The algorithm of Schaul *et al.* [6] exhibits bright artifacts in the grass, trees and bushes. The proposed algorithm outperforms previous methods, preserving color in haze-free regions without compromising the dehazing performance.

Conclusion

This paper studies the intensity inconsistencies, essentially caused by vegetation, between NIR and visible color images. An adaptive algorithm is developed to detect such inconsistencies in a scene and compute a corresponding probabilistic mask. The intensity inconsistencies result in artifacts in such regions for the state-of-the-art dehazing method. Incorporating the inconsistency mask into our image fusion framework, we propose a technique which conserves the photorealism of the original color image during the image fusion. Our results for dehazing a color image using its NIR counterpart are free of artifacts while having state-of-the-art dehazing performance. Future work can study the feasibility of separately dehazing each color channel instead of directly targeting the luma.

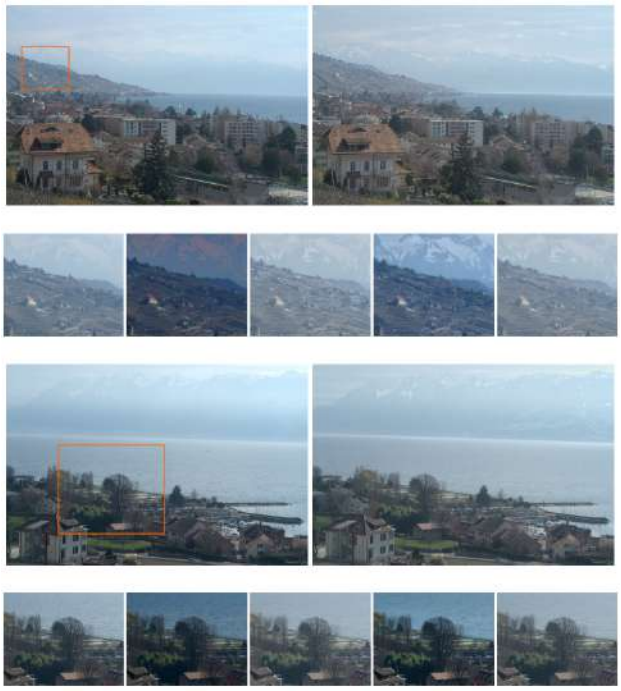


Figure 5. *Original and dehazed images using our approach (full images) and comparison with state-of-the-art results (zooms). From left to right are depicted: original image, He* et al. *[22], Schaul [6], Feng* et al. *[28] and our result. Our method gives state-of-the-art dehazing results while conserving the colors of haze-free regions, thus yielding photorealistic dehazing.*

References

[1] Clément Fredembach and Sabine Süsstrunk, "Colouring the near-infrared," in *Color and Imaging Conference*. Society for Imaging Science and Technology, 2008, pp. 176–182.

[2] Yue M Lu, Clément Fredembach, Martin Vetterli, and Sabine Süsstrunk, "Designing color filter arrays for the joint capture of visible and near-infrared images," in *Proc. IEEE International Conference on Image Processing (ICIP)*, 2009, pp. 3797–3800.

[3] Giacomo Langfelder, Tom Malzbender, A. Longoni, and F. Zaraga, "A device and an algorithm for the separation of visible and near infrared signals in a monolithic silicon sensor," in *IS&T/SPIE Electronic Imaging*, 2011, pp. 788207–788207.

[4] Zahra Sadeghipoor, Yue M Lu, and Sabine Süsstrunk, "A novel compressive sensing approach to simultaneously acquire color and near-infrared images on a single sensor," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 1646–1650.

[5] Xy Luo, J. Zhang, and Q. Dai, "Hybrid fusion and demosaicing algorithm with near-infrared image," in *SPIE Sensing Technology+ Applications*, 2014, pp. 91210J–91210J.

[6] Lex Schaul, Clément Fredembach, and Sabine Süsstrunk, "Color image dehazing using the near-infrared," in *Proc. IEEE International Conference on Image Processing (ICIP)*, 2009, pp. 1629–1632.

[7] Matthew Brown and Sabine Süsstrunk, "Multi-spectral sift for scene category recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 177–184.

[8] Dominic Rüfenacht, Clément Fredembach, and Sabine Süsstrunk,

"Automatic and accurate shadow detection using near-infrared information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 8, pp. 1672–1678, 2014.

[9] Stan Z Li, Rufeng Chu, Shengcai Liao, and Lun Zhang, "Illumination invariant face recognition using near-infrared images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 627–639, 2007.

[10] J.H. Seinfeld and S.N. Pandis, *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, John Wiley and Sons, 2006.

[11] Majed El Helou, Zahra Sadeghipoor, and Sabine Süsstrunk, "Correlation-based near-infrared image deblurring leveraging multispectral chromatic aberration," in *Proc. IEEE International Conference on Image Processing (ICIP)*, 2017.

[12] Jong-Sen Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE transactions on pattern analysis and machine intelligence*, , no. 2, pp. 165–168, 1980.

[13] Yeong-Taeg Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE transactions on Consumer Electronics*, vol. 43, no. 1, pp. 1–8, 1997.

[14] Hui Zhu, Francis H Y Chan, and F K Lam, "Image Contrast Enhancement by Constrained Local Histogram Equalization," *Computer Vision and Image Understanding*, vol. 73, no. 2, pp. 281–290, 1999.

[15] M. Abdullah-Al-Wadud, M. H. Kabir, M. A. Akber Dewan, and O. Chae, "A dynamic histogram equalization for image contrast enhancement," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 593–600, May 2007.

[16] Koen Vande Velde, "Multi-scale color image enhancement," in *International Conference on Image Processing (ICIP)*. IEEE, 1999, vol. 3, pp. 584–587.

[17] Emmanuel J Candès, "Harmonic Analysis of Neural Networks," *Applied and Computational Harmonic Analysis*, vol. 6, no. 2, pp. 197–218, 1999.

[18] J-L Starck, Fionn Murtagh, Emmanuel J Candes, and David L Donoho, "Gray and color image contrast enhancement by the curvelet transform," *IEEE Transactions on image processing*, vol. 12, no. 6, pp. 706–717, 2003.

[19] Robby T. Tan, "Visibility in bad weather from a single image," *26th IEEE Conference on Computer Vision and Pattern Recognition, (CVPR)*, 2008.

[20] Raanan Fattal, "Single image dehazing," *ACM Transactions on Graphics*, vol. 27, no. 3, pp. 1, 2008.

[21] Yoav Y. Schechner, Srinivasa G. Narasimhan, and Shree K. Nayar, "Instant dehazing of images using polarization," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2001, vol. 1, pp. 325–332.

[22] Kaiming He, Jian Sun, and Xiaoou Tang, "Single image haze removal using dark channel prior. Single image haze removal using dark channel prior.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, 2010.

[23] Srinivasa G. Narasimhan and Shree K. Nayar, "Contrast restoration of weather degraded images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 713–724, 2003.

[24] Srinivasa G. Narasimhan and Shree K. Nayar, "Interactive (De) Weathering of an Image using Physical Models," *IEEE Workshop on Color and Photometric Methods in Computer Vision*, pp. 1–8, 2003.

[25] John P Oakley and Brenda L Satherley, "Improving image quality in poor visibility conditions using a physical model for contrast degradation," *IEEE Transactions on Image Processing*, vol. 7, no. 2, pp. 167–179, 1998.

[26] Johannes Kopf, Boris Neubert, Billy Chen, Michael Cohen, Daniel Cohen-Or, Oliver Deussen, Matt Uyttendaele, and Dani Lischinski, "Deep photo: model-based photograph enhancement and viewing," *ACM Trans. Graph.*, vol. 27, no. 5, pp. 1–10, 2008.

[27] Zeev Farbman, Raanan Fattal, Dani Lischinski, and Richard Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," *ACM Transactions on Graphics*, vol. 27, no. 3, pp. 67, 2008.

[28] Chen Feng, Shaojie Zhuo, Xiaopeng Zhang, Liang Shen, and Sabine Ssstrunk, "Near-infrared guided color image dehazing," in *Proc. IEEE International Conference on Image Processing (ICIP)*, 2013, pp. 2363–2367.

[29] David Connah, Mark S. Drew, and Graham D. Finlayson, "Spectral edge: gradient-preserving spectral mapping for image fusion," *Journal of the Optical Society of America A*, vol. 32, no. 12, pp. 2384–2396, 2015.

[30] R. Chouhan and N. Rao, "Vegetation detection in multispectral remote sensing images: Protective role-analysis of vegetation in 2004 indian ocean tsunami," *PDPM Indian Institute of Information Technology*, 2011.

[31] Hui Zou, "The adaptive lasso and its oracle properties," *Journal of the American statistical association*, vol. 101, no. 476, pp. 1418–1429, 2006.

Author Biography

Frederike Dümbgen received her MS in Mechanical Engineering with specialization in robotics from the Ecole Polytechnique Fédérale de Lausanne (2016). She is pursuing a PhD in LCAV under the supervision of Prof. Martin Vetterli and her research interests include indoor localization, simultaneous localization and mapping (SLAM) and computer vision.

Majed El Helou received his BS in Computer and Communications Engineering from the American University of Beirut (2016). He is pursuing his PhD in IVRL at the Ecole Polytechnique Fédérale de Lausanne with Prof. Sabine Süsstrunk. His research interests include signal processing, hyperspectral image processing and artificial intelligence.

Natalija Gucevska received her BS in Computer Science (2016) from the Ecole Polytechnique Fédérale de Lausanne. Since then she is a MS student at Ecole Polytechnique Fédérale de Lausanne in Computer Science specializing in the domain of Images, Signals and Interfaces.